ELSEVIER Linear Algebra and its Applications 279 (1998) 317–324

LINEAR ALGEBRA AND ITS APPLICATIONS

Exponentials of symmetric matrices through tridiagonal reductions

Ya Yan Lu

Department of Mathematics, City University of Hong Kong, Kowloon, Hong Kong, China

Received 10 December 1997; received in revised form 20 January 1998; accepted 29 January 1998

Submitted by R.S. Varga

Abstract

A simple and efficient numerical algorithm for computing the exponential of a symmetric matrix is developed in this paper. For an $n \times n$ matrix, the required number of operations is around $10/3\ n^3$. It is based on the orthogonal reduction to a tridiagonal form and the Chebyshev uniform approximation of e^{-x} on $[0, \infty)$. © 1998 Elsevier Science Inc. All rights reserved.

Keywords: Matrix exponential; Tridiagonal reduction; Chebyshev approximation

1. Introduction

Numerical methods for computing matrix exponentials have been studied by many authors [13,2–4,15,16,19–21]. One popular approach is to use the Padé approximation together with a scaling and squaring strategy. Another method is based on the Schur decomposition and evaluating the exponential of a triangular matrix. The Schur approach is significantly simplified when the matrix is real symmetric or Hermitian. When the spectral decomposition $A = V\Lambda V^T$ is calculated (where A is real symmetric, Λ is the diagonal matrix of the eigenvalues and V is the orthogonal matrix of the eigenvectors), we simply evaluate e^A by $Ve^{\Lambda}V^T$. The spectral decomposition $A = V\Lambda V^T$ requires about $9\ n^3$ (n is the size of the matrix) operations by the standard QR method and it requires about $4\ n^3$ operations if Cuppen's divide-and-conquer method [7] is used. High quality implementations of these methods are available in LAPACK [1]. The matrix multiplications in $Ve^{\Lambda}V^T$ require another n^3 operations, since e^A is also

PII: S0024-3795(98)00017-2

symmetric. Therefore, if the Schur approach is used for symmetric matrices, at least 5 n^3 operations are required. The symmetry of the matrix A can also be used in the Padé method to save some operations. However, the scaling and squaring procedure depends on the norm of A. If A has a large norm, more squarings are needed leading to larger round off errors and more required operations.

For a symmetric matrix, the orthogonal reduction to a tridiagonal matrix is the first step in computing its eigenvalues and eigenvectors. This gives rise to the decomposition $A = QTQ^{\mathrm{T}}$, where Q is orthogonal and T is symmetric tridiagonal. For evaluating matrix functions, the formula $f(A) = Qf(T)Q^{\mathrm{T}}$ may be useful. For the square root and logarithm functions, this approach has lead to efficient numerical methods for symmetric positive definite matrices [11,12]. When A is full, the reduction $A = QTQ^{\mathrm{T}}$ requires $4/3\ n^3$ operations using Householder reflections. After $f(T)$ is calculated, the sequence of these reflections are applied back to $f(T)$. This last step requires 2 n^3 operations. Therefore, the total number of required operations for the tridiagonal reduction approach is at least $10/3\ n^3$. In this paper, we calculate $f(T)$ based on the Chebyshev uniform approximation for e^{-x} on $[0, \infty)$. The required number of operations for $f(T)$ is only $\mathrm{O}(n^2)$. The leading term of the total number of required operations for computing e^A is thus still $10/3\ n^3$.

In Section 2, we discuss the details of our method. Numerical experiments are reported in Section 3.

2. The new method

The rational Chebyshev approximation of e^{-x} on $[0, \infty)$ was first studied by Varga [17] in 1961. For a nonnegative integer p, a rational function $R_p(x) = N_p(x)/D_p(x)$ (where N_p and D_p are polynomials of x with degree p) can be found with the smallest maximum error $\max_{0 \leqslant x < \infty} |R_p(x) - \mathrm{e}^{-x}| = \tau_{p,p}$, where $\tau_{p,p}$ is the uniform rational Chebyshev constant [18]. Cody, Meinardus and Varga [6] proved that $\tau_{p,p}$ converges geometrically to zero. The exact geometric convergence factor was determined by Gonchar and Rakhmanov [10]. This gives rise to

$$\lim_{p \to \infty} \tau_{p,p}^{1/p} = \frac{1}{9.2890\ldots}.$$

The coefficients for R_p are listed in [5] for $2 \leqslant p \leqslant 30$. For our purpose here, we need the partial fraction expansion

$$R_p(x) = \alpha_0 + \sum_{j=1}^{p} \frac{\alpha_j}{x - \theta_j}.$$

The poles $\{\theta_j\}$ and the coefficients $\{\alpha_j\}$ were computed and listed in Gallopoulos and Saad [8] for $p = 10$ and $p = 14$. If T is a positive semi-definite matrix, we have

$$e^{-T} \approx \alpha_0 I + \sum_{j=1}^{p} \alpha_j (T - \theta_j I)^{-1}.$$

The error of the above approximation is bounded by $\tau_{p,p}$ in the matrix 2-norm. From [6,5], we have $\tau_{14,14} \approx 1.832174 \times 10^{-14}$. Therefore, we should expect 14 digits of accuracy when $p = 14$ is used. Since these coefficients and poles appears in complex conjugate pairs, it is sufficient to compute the real parts of seven terms for $p = 14$. In general, the matrix T is not positive semi-definite. However, we can easily calculate the smallest eigenvalue (say, λ_1) of T and use the Chebyshev rational approximation for $T - \lambda_1 I$.

Our algorithm for computing the (negative) exponential of a symmetric matrix takes the following steps:

1. Reduce the matrix A to a tridiagonal matrix T by orthogonal similarity transformations, i.e., $A = QTQ^{\mathrm{T}}$, where Q is an orthogonal matrix.
2. Calculate the smallest eigenvalue of T, say λ_1.
3. Approximate $\exp(-(T - \lambda_1 I))$ by

   $$R_p = \alpha_0 I + \sum_{j=1}^{p} \alpha_j [T - (\lambda_1 + \theta_j) I]^{-1}.$$

   Typically, we choose $p = 14$ and use $\{\alpha_j, \theta_j\}$ given in [8].
4. Evaluate the approximation of $\exp(-A)$ by

   $$e^{-A} \approx S = e^{-\lambda_1} Q R_p Q^{\mathrm{T}}.$$

Step 1 is the standard first step used in calculating eigenvalues and eigenvectors of symmetric matrices [9]. Householder reflections or Givens rotations (for banded matrices) can be used to compute the decomposition. The orthogonal matrix Q is not explicitly formed. It is implicitly given as a sequence of the reflections or rotations. For a full matrix A, the required number of operations is $4/3\ n^3$. In our implementation, we use the LAPACK routine xSYTRD. In Step 2, we calculate the smallest eigenvalue of the symmetric tridiagonal matrix T. The required number of operations is $O(n)$ if the bisection method is used. We use the LAPACK routine xSTEBZ in our program.

The inverse of a tridiagonal matrix can be found in $O(n^2)$ operations. In Step 3, we use LU decomposition with partial pivoting for the complex tridiagonal matrices, then solve the columns of the inverse. A further reduction in the required number of operations is possible, if we take advantage of the symmetry and the zeros in the identity matrix (as the right-hand side in the equation for the inverse). The matrix $T - (\lambda_1 + \theta_j)I$ can not be near singular, since T and λ_1 are real and θ_j has a non-zero imaginary part. For $p = 14$, the imaginary parts of θ_j range from 1.19 to 16.6. The coefficients $\{\alpha_j\}$ are of moderate size. There is no serious loss of significance when the inverse matrices are added together.

Step 4 involves the multiplication of Householder reflections (or Givens rotations) back on R_p from both the left and right sides. This is the opposite of the tridiagonal reduction in Step 1. Similar techniques as in Step 1 can be used here to reduce the computational complexity. However, this step is more expensive, since it does not produces zeros. The required number of operations for this step is around 2 n^3. The multiplication of the scalar $e^{-\lambda_1}$ can be avoided, if it is first multiplied to $\{\alpha_j\}$.

The total required number of operations for computing $\exp(-A)$ is around $10/3\ n^3$, and the main contributions are from Steps 1 and 4.

Theoretically, since the rational Chebyshev approximation has the maximum error $\tau_{p,p}$, we have

$$||e^{-A} - S||_2 \leqslant e^{-\lambda_1}\tau_{p,p} = ||e^{-A}||_2\tau_{p,p},$$

where S is the approximation of $\exp(-A)$ in Step 4.

3. Numerical examples

Our program for symmetric matrix exponentials is implemented in double precision based on $\{\theta_j, \alpha_j\}$ for $p = 14$ given in [8]. These values of θ_j and α_j are not as accurate as we would like. In fact, the maximum error based on these approximate values is 3.13×10^{-12} at $x = 0$. Therefore, we expect to have 12 correct digits. To verify the accuracy of our computations, we need the "exact" solution to compare with. For this purpose, we developed a quadruple precision routine that solves the symmetric matrix eigenvalue problem ($A = V\Lambda V^{\mathrm{T}}$) and evaluates e^{-A} by $Ve^{-\Lambda}V^{\mathrm{T}}$. This is based on a simple modification of the EISPACK routines (`TRED2`, `IMTQL2` and `PYTHAG`) for symmetric matrix eigenvalue problems. The accuracy of the modified programs is independently verified by a multi-precision calculation in MAPLE. The numerical solution obtained from the quadruple precision routine is then used as the "exact solution" for comparison. We calculate the relative errors in various matrix norms:

$$e_f = \frac{||e^{-A} - S||_f}{||e^{-A}||_f}, \quad e_1 = \frac{||e^{-A} - S||_1}{||e^{-A}||_1}, \quad e_2 = \frac{||e^{-A} - S||_2}{||e^{-A}||_2},$$

where S is the numerical approximation of e^{-A} as in Step 4.

Example 1. The (i,j) entry of the $n \times n$ matrix A is

$$a_{ij} = \frac{1}{2 + (i-j)^2}.$$

The relative errors in different matrix norms are listed in Table 1.

Table 1
Relative errors in computing $\exp(-A)$ for Example 1

n	e_f	e_1	e_2
4	2.83×10^{-12}	3.21×10^{-12}	3.13×10^{-12}
10	2.82×10^{-12}	2.99×10^{-12}	3.13×10^{-12}
20	2.83×10^{-12}	2.94×10^{-12}	3.13×10^{-12}
40	2.84×10^{-12}	2.92×10^{-12}	3.13×10^{-12}
60	2.84×10^{-12}	2.92×10^{-12}	3.13×10^{-12}
80	2.84×10^{-12}	2.92×10^{-12}	3.13×10^{-12}
100	2.84×10^{-12}	2.92×10^{-12}	3.13×10^{-12}
200	2.84×10^{-12}	2.93×10^{-12}	3.13×10^{-12}

Example 2. The $n \times n$ matrix A is obtained from a second order finite difference discretization of the negative Laplacian on a unit square with Dirichlet boundary conditions. The (i,j) entry is given by

$$a_{ij} = \begin{cases} 4 & \text{if } i = j, \\ -1 & \text{if } |i-j| = p, \\ -1 & \text{if } |i-j| = 1 \text{ and } (i+j) \bmod (2p) \neq 1, \\ 0 & \text{otherwise,} \end{cases}$$

where $n = p^2$. The results are listed in Table 2.

Example 3. We consider symmetric matrices whose entries are random numbers from the uniform distribution on the interval $[-1/2, 1/2]$. In Table 3, we list the results for ten 100×100 symmetric matrices.

These results are very satisfactory. The relative errors in the matrix 2-norm are always bounded by 3.13×10^{-12}. Relative errors in Frobenius norm and 1-norm are also at the same order. It seems safe to say that 12 correct digits are obtained.

Our method is also very efficient compared with other methods. In Table 4, we list the required execution time in seconds for our method, the Padé method (with scaling and squaring) and the Schur method (which solves the eigenvalue problem first). For the Padé method, we use the implementation by Sidje [14]. Our implementation of the Schur method (only for symmetric matrices) is based on the LAPACK implementation of Cuppen's divide-and-conquer [7] method for symmetric matrix eigenvalue problems and a straight forward implementation of $Ve^{-\Lambda}V^{\mathrm{T}}$ which requires about n^3 operations. All programs are in double precision and compiled (including LAPACK) with the same option-`fast` on a SUN Ultra 1 (model 170) workstation with the compiler `f77` version 4.0 from SUN Microsystems.

In Table 4, if we implement the Schur method based on the QR algorithm for symmetric matrix eigenvalue problem, the required execution time

Table 2
Relative errors in computing $\exp(-A)$ for Example 2

n	e_f	e_1	e_2
4	3.08×10^{-12}	3.13×10^{-12}	3.13×10^{-12}
9	2.98×10^{-12}	3.13×10^{-12}	3.13×10^{-12}
16	2.83×10^{-12}	3.21×10^{-12}	3.13×10^{-12}
25	2.70×10^{-12}	3.42×10^{-12}	3.13×10^{-12}
36	2.60×10^{-12}	3.45×10^{-12}	3.13×10^{-12}
49	2.52×10^{-12}	3.57×10^{-12}	3.13×10^{-12}
64	2.47×10^{-12}	3.50×10^{-12}	3.13×10^{-12}
81	2.45×10^{-12}	3.49×10^{-12}	3.13×10^{-12}
100	2.43×10^{-12}	3.44×10^{-12}	3.13×10^{-12}
121	2.41×10^{-12}	3.42×10^{-12}	3.13×10^{-12}
144	2.39×10^{-12}	3.40×10^{-12}	3.13×10^{-12}
169	2.38×10^{-12}	3.38×10^{-12}	3.13×10^{-12}
196	2.37×10^{-12}	3.37×10^{-12}	3.13×10^{-12}

Table 3
Relative errors in computing $\exp(-A)$ for 100×100 random matrices

n	e_f	e_1	e_2
100	2.44×10^{-12}	2.98×10^{-12}	3.13×10^{-12}
100	2.49×10^{-12}	2.92×10^{-12}	3.13×10^{-12}
100	2.43×10^{-12}	2.80×10^{-12}	3.13×10^{-12}
100	2.49×10^{-12}	2.97×10^{-12}	3.13×10^{-12}
100	2.43×10^{-12}	3.00×10^{-12}	3.13×10^{-12}
100	2.58×10^{-12}	2.79×10^{-12}	3.13×10^{-12}
100	2.43×10^{-12}	2.78×10^{-12}	3.13×10^{-12}
100	2.49×10^{-12}	3.04×10^{-12}	3.13×10^{-12}
100	2.47×10^{-12}	2.98×10^{-12}	3.13×10^{-12}
100	2.46×10^{-12}	2.71×10^{-12}	3.13×10^{-12}

is longer. Our method has a clear advantage for large n ($n = 300$ and 400 as in the table), since the required number of operations has the smallest leading order term.

4. Conclusion

The method developed in this paper relies on the orthogonal reduction of a symmetric matrix to the tridiagonal form and the uniform Chebyshev rational approximation of e^{-x} on $[0, +\infty)$. The rational approximation is given in the prime fraction form and it is applied to the shifted (semi-definite) tridiagonal matrix. Theoretically, the relative error in the matrix 2-norm for

Table 4
Time (in seconds) required for computing e^{-A} by different methods

Matrix	n	T_{new}	T_{pade}	T_{schur}
Example 1	100	0.15	0.24	0.13
Example 1	200	0.67	3.09	0.81
Example 1	300	1.99	13.7	3.40
Example 1	400	5.39	36.4	8.40
Example 2	100	0.15	0.15	0.10
Example 2	196	0.64	1.89	0.55
Example 2	324	2.47	9.73	3.54
Example 2	400	4.83	19.3	6.08
Example 3	100	0.15	0.30	0.14
Example 3	200	0.69	4.82	0.90
Example 3	300	2.00	20.1	3.50
Example 3	400	5.05	52.8	9.27

approximating the (symmetric) matrix exponential is bounded by the uniform rational Chebyshev constant $\tau_{p,p}$. Unlike the Padé method with scaling and squaring, the accuracy of our solution is not related to the norm of the matrix A. The method requires about $10/3\ n^3$ operations and it is more efficient than the Padé method and the Schur decomposition method.

References

[1] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, D. Sorensen, LAPACK User's Guide, SIAM, 1992.
[2] M. Arioli, B. Codenotti, C. Fassino, The Padé method for computing the matrix exponential, Linear Algebra Appl. 240 (1996) 111–130.
[3] R. Bellman, On the calculation of matrix exponential, Linear and Multilinear Algebra 13 (1) (1983) 73–79.
[4] P. Bochev, S. Markov, A self-validating numerical method for the matrix exponential, Computing 43 (1) (1989) 59–72.
[5] A.J. Carpenter, A. Ruttan, R.S. Varga, Extended numerical computations on the 1/9 conjecture in rational approximation theory, Lecture Notes in Mathematics 1105, Springer, Berlin, 1984, pp. 383–411.
[6] W.J. Cody, G. Meinardus, R.S. Varga, Chebyshev rational approximation to $\exp(-x)$ in $[0, +\infty)$ and applications to heat conduction problems, J. Approx. Theory 2 (1969) 50–65.
[7] J.J.M. Cuppen, A divide and conquer method for the symmetric eigenproblem, Numer. Math. 36 (1981) 177–195.
[8] G. Gallopoulos, Y. Saad, Efficient solution of parabolic equations by Krylov approximation methods, SIAM J. Sci. Statist. Comput. 13 (1992) 1236–1264.
[9] G.H. Golub, C.F. Van Loan, Matrix Computations, 2nd ed., The Johns Hopkins University Press, Baltimore, MD, 1989.

[10] A.A. Gonchar, E.A. Rakhmanov, Equilibrium distribution and the degree of rational approximation of analytic functions, Mat. Sbornik 134 (176) (1987) 306–352 (in Russian); English translation in Math. USSR Sbnornik 62 (1989) 305–348.
[11] Y.Y. Lu, A Padé approximation method for square roots of symmetric positive definite matrices, SIAM Journal on Matrix Analysis and Applications, accepted for publication.
[12] Y.Y. Lu, Computing the logarithm of a symmetric positive definite matrix, Applied Numerical Mathematics, accepted for publication.
[13] C.B. Moler, C.F. Van Loan, Nineteen dubious ways to compute the exponential of a matrix, SIAM Rev. 20 (1978) 801–836.
[14] R.B. Sidje, EXPOKIT: Software package for computing matrix exponentials, to appear in ACM – Transactions on Mathematical Software.
[15] E.U. Stickel, A splitting method for the calculation of the matrix exponential, Analysis 14 (2) (1994) 103–112.
[16] C. Van Loan, The sensitivity of the matrix exponential, SIAM J. Numer. Anal. 14 (6) (1977) 971–981.
[17] R.S. Varga, On higher order stable implicit methods for solving parabolic partial differential equations, J. Math. and Phys. XL (1961) 220–231.
[18] R.S. Varga, Scientific computation on mathematical problems and conjectures, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, 60, 1990.
[19] R.C. Ward, Numerical computation of the matrix exponential with accuracy estimate, SIAM J. Numer. Anal. 14 (4) (1977) 600–610.
[20] G. Walz, Computing the matrix exponential and other matrix functions, J. Comput. Appl. Math. 21 (1) (1988) 119–123.
[21] D.W. Yu, On the numerical computation of the matrix exponential, J. Korean Math. Soc. 31 (4) (1994) 633–643.